



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: SHYJAN, Andrew
- (ii) TITLE OF INVENTION: NOVEL MULTIDRUG RESISTANCE-ASSOCIATED POLYPEPTIDE
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: LAHIVE & COCKFIELD, LLP
(B) STREET: 28 State Street
(C) CITY: Boston
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02109
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 09/061,400
(B) FILING DATE: 16-APRIL-1998
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Elizabeth A. Hanley
(B) REGISTRATION NUMBER: 33,505
(C) REFERENCE/DOCKET NUMBER: MNI-056CP
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 227-7400
(B) TELEFAX: (617) 742-4214

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4847 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 116..4426
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- | | |
|---|-----|
| GGCTCATGCT CGGGAGCGTG GTTGAGCGGC TGGCGCGGTT GTCCTGGAGC AGGGGCGCAG | 60 |
| GAATTCTGAT GTGAAACTAA CAGTCTGTGA GCCCTGGAAC CTCCACTCAG AGAAG ATG | 118 |
| Met | |
| 1 | |
| AAG GAT ATC GAC ATA GGA AAA GAG TAT ATC ATC CCC AGT CCT GGG TAT | 166 |
| Lys Asp Ile Asp Ile Gly Lys Glu Tyr Ile Ile Pro Ser Pro Gly Tyr | |
| 5 10 15 | |
| AGA AGT GTG AGG GAG AGA ACC AGC ACT TCT GGG ACG CAC AGA GAC CGT | 214 |
| Arg Ser Val Arg Glu Arg Thr Ser Thr Ser Gly Thr His Arg Asp Arg | |
| 20 25 30 | |
| GAA GAT TCC AAG TTC AGG AGA ACT CGA CCG TTG GAA TGC CAA GAT GCC | 262 |
| Glu Asp Ser Lys Phe Arg Arg Thr Arg Pro Leu Glu Cys Gln Asp Ala | |

35	40	45	
TTG GAA ACA GCA GCC CGA GCC GAG GGC CTC TCT CTT GAT GCC TCC ATG Leu Glu Thr Ala Ala Arg Ala Glu Gly Leu Ser Leu Asp Ala Ser Met 50 55 60 65			310
CAT TCT CAG CTC AGA ATC CTG GAT GAG GAG CAT CCC AAG GGA AAG TAC His Ser Gln Leu Arg Ile Leu Asp Glu Glu His Pro Lys Gly Lys Tyr 70 75 80			358
CAT CAT GGC TTG AGT GCT CTG AAG CCC ATC CGG ACT ACT TGC AAA CAC His His Gly Leu Ser Ala Leu Lys Pro Ile Arg Thr Thr Cys Lys His 85 90 95			406
CAG CAC CCA GTG GAC AAT GCT GGG CTT TTT TCC TGT ATG ACT TTT TCG Gln His Pro Val Asp Asn Ala Gly Leu Phe Ser Cys Met Thr Phe Ser 100 105 110			454
TGG CTT TCT TCT CTG GCC CGT GTG GCC CAC AAG AAG GGG GAG CTC TCA Trp Leu Ser Ser Leu Ala Arg Val Ala His Lys Lys Gly Glu Leu Ser 115 120 125			502
ATG GAA GAC GTG TGG TCT CTG TCC AAG CAC GAG TCT TCT GAC GTG AAC Met Glu Asp Val Trp Ser Leu Ser Lys His Glu Ser Ser Asp Val Asn 130 135 140 145			550
TGC AGA AGA CTA GAG AGA CTG TGG CAA GAA GAG CTG AAT GAA GTT GGG Cys Arg Arg Leu Glu Arg Leu Trp Gln Glu Glu Leu Asn Glu Val Gly 150 155 160			598
CCA GAC GCT GCT TCC CTG CGA AGG GTT GTG TGG ATC TTC TGC CGC ACC Pro Asp Ala Ser Leu Arg Arg Val Val Trp Ile Phe Cys Arg Thr 165 170 175			646
AGG CTC ATC CTG TCC ATC GTG TGC CTG ATG ATC ACG CAG CTG GCT GGC Arg Leu Ile Leu Ser Ile Val Cys Leu Met Ile Thr Gln Leu Ala Gly 180 185 190			694
TTC AGT GGA CCA GCC TTC ATG GTG AAA CAC CTC TTG GAG TAT ACC CAG Phe Ser Gly Pro Ala Phe Met Val Lys His Leu Leu Glu Tyr Thr Gln 195 200 205			742
GCA ACA GAG TCT AAC CTG CAG TAC AGC TTG TTG TTA GTG CTG GGC CTC Ala Thr Glu Ser Asn Leu Gln Tyr Ser Leu Leu Leu Val Leu Gly Leu 210 215 220 225			790
CTC CTG ACG GAA ATC GTG CGG TCT TGG TCG CTT GCA CTG ACT TGG GCA Leu Leu Thr Glu Ile Val Arg Ser Trp Ser Leu Ala Leu Thr Trp Ala 230 235 240			838
TTG AAT TAC CGA ACC GGT GTC CGC TTG CGG GGG GCC ATC CTA ACC ATG Leu Asn Tyr Arg Thr Gly Val Arg Leu Arg Gly Ala Ile Leu Thr Met 245 250 255			886
GCA TTT AAG AAG ATC CTT AAG TTA AAG AAC ATT AAA GAG AAA TCC CTG Ala Phe Lys Lys Ile Leu Lys Leu Lys Asn Ile Lys Glu Lys Ser Leu 260 265 270			934
GGT GAG CTC ATC AAC ATT TGC TCC AAC GAT GGG CAG AGA ATG TTT GAG Gly Glu Leu Ile Asn Ile Cys Ser Asn Asp Gly Gln Arg Met Phe Glu 275 280 285			982
GCA GCA GCC GTT GGC AGC CTG CTG GCT GGA GGA CCC GTT GTT GCC ATC Ala Ala Ala Val Gly Ser Leu Leu Ala Gly Pro Val Val Ala Ile 290 295 300 305			1030
TTA GGC ATG ATT TAT AAT GTA ATT ATT CTG GGA CCA ACA GGC TTC CTG Leu Gly Met Ile Tyr Asn Val Ile Ile Leu Gly Pro Thr Gly Phe Leu 310 315 320			1078
GGA TCA GCT GTT TTT ATC CTC TTT TAC CCA GCA ATG ATG TTT GCA TCA Gly Ser Ala Val Phe Ile Leu Phe Tyr Pro Ala Met Met Phe Ala Ser			1126

B1
Am4

325					330					335						
CGG Arg	CTC Leu	ACA Thr 340	GCA Ala	TAT Tyr	TTC Phe	AGG Arg	AGA Arg 345	AAA Lys	TGC Cys	GTG Val	GCC Ala	GCC Ala 350	ACG Thr	GAT Asp	GAA Glu	1174
CGT Arg	GTC Val 355	CAG Gln	AAG Lys	ATG Met	AAT Asn	GAA Glu 360	GTT Val	CTT Leu	ACT Thr	TAC Tyr	ATT Ile 365	AAA Lys	TTT Phe	ATC Ile	AAA Lys	1222
ATG Met 370	TAT Tyr	GCC Ala	TGG Trp	GTC Val	AAA Lys 375	GCA Ala	TTT Phe	TCT Ser	CAG Gln	AGT Ser 380	GTT Val	CAG Gln	AAA Lys	ATC Ile	CGC Arg 385	1270
GAG Glu	GAG Glu	GAG Glu	CGT Arg	CGG Arg 390	ATA Ile	TTG Leu	GAA Glu	AAA Lys	GCC Ala 395	GGG Gly	TAC Tyr	TTC Phe	CAG Gln	AGC Ser 400	ATC Ile	1318
ACT Thr	GTG Val	GGT Gly 405	GTG Val	GCT Ala	CCC Pro	ATT Ile	GTG Val 410	GTG Val	GTG Val	ATT Ile	GCC Ala	AGC Ser	GTG Val 415	GTG Val	ACC Thr	1366
TTC Phe	TCT Ser	GTT Val 420	CAT His	ATG Met	ACC Thr	CTG Leu	GGC Gly 425	TTC Phe	GAT Asp	CTG Leu	ACA Thr	GCA Ala 430	GCA Ala	CAG Gln	GCT Ala	1414
TTC Phe	ACA Thr 435	GTG Val	GTG Val	ACA Thr	GTC Val	TTC Phe 440	AAT Asn	TCC Ser	ATG Met	ACT Thr	TTT Phe 445	GCT Ala	TTG Leu	AAA Lys	GTA Val	1462
ACA Thr 450	CCG Pro	TTT Phe	TCA Ser	GTA Val	AAG Lys 455	TCC Ser	CTC Leu	TCA Ser	GAA Glu	GCC Ala 460	TCA Ser	GTG Val	GCT Ala	GTT Val	GAC Asp 465	1510
AGA Arg	TTT Phe	AAG Lys	AGT Ser	TTG Leu 470	TTT Phe	CTA Leu	ATG Met	GAA Glu	GAG Glu 475	GTT Val	CAC His	ATG Met	ATA Ile	AAG Lys 480	AAC Asn	1558
AAA Lys	CCA Pro	GCC Ala	AGT Ser 485	CCT Pro	CAC His	ATC Ile	AAG Lys	ATA Ile 490	GAG Glu	ATG Met	AAA Lys	AAT Asn	GCC Ala 495	ACC Thr	TTG Leu	1606
GCA Ala	TGG Trp	GAC Asp 500	TCC Ser	TCC Ser	CAC His	TCC Ser	AGT Ser 505	ATC Ile	CAG Gln	AAC Asn	TCG Ser	CCC Pro 510	AAG Lys	CTG Leu	ACC Thr	1654
CCC Pro	AAA Lys 515	ATG Met	AAA Lys	AAA Lys	GAC Asp	AAG Lys 520	AGG Arg	GCT Ala	TCC Ser	AGG Arg	GGC Gly 525	AAG Lys	AAA Lys	GAG Glu	AAG Lys	1702
GTG Val 530	AGG Arg	CAG Gln	CTG Leu	CAG Gln	CGC Arg	ACT Thr	GAG Glu	CAT His	CAG Gln	GCG Ala 540	GTG Val	CTG Leu	GCA Ala	GAG Glu	CAG Gln 545	1750
AAA Lys	GGC Gly	CAC His	CTC Leu 550	CTC Leu	CTG Leu	GAC Asp	AGT Ser	GAC Asp	GAG Glu 555	CGG Arg	CCC Pro	AGT Ser	CCC Pro	GAA Glu 560	GAG Glu	1798
GAA Glu	GAA Glu	GGC Gly	AAG Lys 565	CAC His	ATC Ile	CAC His	CTG Leu	GGC Gly 570	CAC His	CTG Leu	CGC Arg	TTA Leu	CAG Gln 575	AGG Arg	ACA Thr	1846
CTG Leu	CAC His	AGC Ser 580	ATC Ile	GAT Asp	CTG Leu	GAG Glu	ATC Ile 585	CAA Gln	GAG Glu	GGT Gly	AAA Lys	CTG Leu 590	GTT Val	GGA Gly	ATC Ile	1894
TGC Cys	GGC Gly 595	AGT Ser	GTG Val	GGA Gly	AGT Ser	GGA Gly 600	AAA Lys	ACC Thr	TCT Ser	CTC Leu	ATT Ile 605	TCA Ser	GCC Ala	ATT Ile	TTA Leu	1942
GGC Gly	CAG Gln	ATG Met	ACG Thr	CTT Leu	CTA Leu	GAG Glu	GGC Gly	AGC Ser	ATT Ile	GCA Ala	ATC Ile	AGT Ser	GGA Gly	ACC Thr	TTC Phe	1990

610	615	620	625	
GCT TAT GTG GCC CAG CAG GCC TGG ATC CTC AAT GCT ACT CTG AGA GAC Ala Tyr Val Ala Gln Gln Ala Trp Ile Leu Asn Ala Thr Leu Arg Asp 630 635 640				2038
AAC ATC CTG TTT GGG AAG GAA TAT GAT GAA GAA AGA TAC AAC TCT GTG Asn Ile Leu Phe Gly Lys Glu Tyr Asp Glu Glu Arg Tyr Asn Ser Val 645 650 655				2086
CTG AAC AGC TGC TGC CTG AGG CCT GAC CTG GCC ATT CTT CCC AGC AGC Leu Asn Ser Cys Cys Leu Arg Pro Asp Leu Ala Ile Leu Pro Ser Ser 660 665 670				2134
GAC CTG ACG GAG ATT GGA GAG CGA GGA GCC AAC CTG AGC GGT GGG CAG Asp Leu Thr Glu Ile Gly Glu Arg Gly Ala Asn Leu Ser Gly Gly Gln 675 680 685				2182
CGC CAG AGG ATC AGC CTT GCC CGG GCC TTG TAT AGT GAC AGG AGC ATC Arg Gln Arg Ile Ser Leu Ala Arg Ala Leu Tyr Ser Asp Arg Ser Ile 690 695 700 705				2230
TAC ATC CTG GAC GAC CCC CTC AGT GCC TTA GAT GCC CAT GTG GGC AAC Tyr Ile Leu Asp Asp Pro Leu Ser Ala Leu Asp Ala His Val Gly Asn 710 715 720				2278
CAC ATC TTC AAT AGT GCT ATC CGG AAA CAT CTC AAG TCC AAG ACA GTT His Ile Phe Asn Ser Ala Ile Arg Lys His Leu Lys Ser Lys Thr Val 725 730 735				2326
CTG TTT GTT ACC CAC CAG TTA CAG TAC CTG GTT GAC TGT GAT GAA GTG Leu Phe Val Thr His Gln Leu Tyr Leu Val Asp Cys Asp Glu Val 740 745 750				2374
ATC TTC ATG AAA GAG GGC TGT ATT ACG GAA AGA GGC ACC CAT GAG GAA Ile Phe Met Lys Glu Gly Cys Ile Thr Glu Arg Gly Thr His Glu Glu 755 760 765				2422
CTG ATG AAT TTA AAT GGT GAC TAT GCT ACC ATT TTT AAT AAC CTG TTG Leu Met Asn Leu Asn Gly Asp Tyr Ala Thr Ile Phe Asn Asn Leu Leu 770 775 780 785				2470
CTG GGA GAG ACA CCG CCA GTT GAG ATC AAT TCA AAA AAG GAA ACC AGT Leu Gly Glu Thr Pro Pro Val Glu Ile Asn Ser Lys Lys Glu Thr Ser 790 795 800				2518
GGT TCA CAG AAG AAG TCA CAA GAC AAG GGT CCT AAA ACA GGA TCA ATA Gly Ser Gln Lys Lys Ser Gln Asp Lys Gly Pro Lys Thr Gly Ser Ile 805 810 815				2566
AAG AAG GAA AAA GCA GTA AAG CCA GAG GAA GGG CAG CTT GTG CAG CTG Lys Lys Glu Lys Ala Val Lys Pro Glu Glu Gly Gln Leu Val Gln Leu 820 825 830				2614
GAA GAG AAA GGG CAG GGT TCA GTG CCC TGG TCA GTA TAT GGT GTC TAC Glu Glu Lys Gly Gln Gly Ser Val Pro Trp Ser Val Tyr Gly Val Tyr 835 840 845				2662
ATC CAG GCT GCT GGG GGC CCC TTG GCA TTC CTG GTT ATT ATG GCC CTT Ile Gln Ala Ala Gly Gly Pro Leu Ala Phe Leu Val Ile Met Ala Leu 850 855 860 865				2710
TTC ATG CTG AAT GTA GGC AGC ACC GCC TTC AGC ACC TGG TGG TTG AGT Phe Met Leu Asn Val Gly Ser Thr Ala Phe Ser Thr Trp Trp Leu Ser 870 875 880				2758
TAC TGG ATC AAG CAA GGA AGC GGG AAC ACC ACT GTG ACT CGA GGG AAC Tyr Trp Ile Lys Gln Gly Ser Gly Asn Thr Thr Val Thr Arg Gly Asn 885 890 895				2806
GAG ACC TCG GTG AGT GAC AGC ATG AAG GAC AAT CCT CAT ATG CAG TAC Glu Thr Ser Val Ser Asp Ser Met Lys Asp Asn Pro His Met Gln Tyr				2854

B1
Gmt

900					905					910						
TAT	GCC	AGC	ATC	TAC	GCC	CTC	TCC	ATG	GCA	GTC	ATG	CTG	ATC	CTG	AAA	2902
Tyr	Ala	Ser	Ile	Tyr	Ala	Leu	Ser	Met	Ala	Val	Met	Leu	Ile	Leu	Lys	
915					920					925						
GCC	ATT	CGA	GGA	GTT	GTC	TTT	GTC	AAG	GGC	ACG	CTG	CGA	GCT	TCC	TCC	2950
Ala	Ile	Arg	Gly	Val	Val	Phe	Val	Lys	Gly	Thr	Leu	Arg	Ala	Ser	Ser	
930					935					940					945	
CGG	CTG	CAT	GAC	GAG	CTT	TTC	CGA	AGG	ATC	CTT	CGA	AGC	CCT	ATG	AAG	2998
Arg	Leu	His	Asp	Glu	Leu	Phe	Arg	Arg	Ile	Leu	Arg	Ser	Pro	Met	Lys	
950					955					960						
TTT	TTT	GAC	ACG	ACC	CCC	ACA	GGG	AGG	ATT	CTC	AAC	AGG	TTT	TCC	AAA	3046
Phe	Phe	Asp	Thr	Thr	Pro	Thr	Gly	Arg	Ile	Leu	Asn	Arg	Phe	Ser	Lys	
965					970					975						
GAC	ATG	GAT	GAA	GTT	GAC	GTG	CGG	CTG	CCG	TTC	CAG	GCC	GAG	ATG	TTC	3094
Asp	Met	Asp	Glu	Val	Asp	Val	Arg	Leu	Pro	Phe	Gln	Ala	Glu	Met	Phe	
980					985					990						
ATC	CAG	AAC	GTT	ATC	CTG	GTG	TTC	TTC	TGT	GTG	GGA	ATG	ATC	GCA	GGA	3142
Ile	Gln	Asn	Val	Ile	Leu	Val	Phe	Phe	Cys	Val	Gly	Met	Ile	Ala	Gly	
995					1000					1005						
GTC	TTC	CCG	TGG	TTC	CTT	GTG	GCA	GTG	GGG	CCC	CTT	GTC	ATC	CTC	TTT	3190
Val	Phe	Pro	Trp	Phe	Leu	Val	Ala	Val	Gly	Pro	Leu	Val	Ile	Leu	Phe	
1010					1015					1020					1025	
TCA	GTC	CTG	CAC	ATT	GTC	TCC	AGG	GTC	CTG	ATT	CGG	GAG	CTG	AAG	CGT	3238
Ser	Val	Leu	His	Ile	Val	Ser	Arg	Val	Leu	Ile	Arg	Glu	Leu	Lys	Arg	
1030					1035					1040						
CTG	GAC	AAT	ATC	ACG	CAG	TCA	CCT	TTC	CTC	TCC	CAC	ATC	ACG	TCC	AGC	3286
Leu	Asp	Asn	Ile	Thr	Gln	Ser	Pro	Phe	Leu	Ser	His	Ile	Thr	Ser	Ser	
1045					1050					1055						
ATA	CAG	GGC	CTT	GCC	ACC	ATC	CAC	GCC	TAC	AAT	AAA	GGG	CAG	GAG	TTT	3334
Ile	Gln	Gly	Leu	Ala	Thr	Ile	His	Ala	Tyr	Asn	Lys	Gly	Gln	Glu	Phe	
1060					1065					1070						
CTG	CAC	AGA	TAC	CAG	GAG	CTG	CTG	GAT	GAC	AAC	CAA	GCT	CCT	TTT	TTT	3382
Leu	His	Arg	Tyr	Gln	Glu	Leu	Leu	Asp	Asp	Asn	Gln	Ala	Pro	Phe	Phe	
1075					1080					1085						
TTG	TTT	ACG	TGT	GCG	ATG	CGG	TGG	CTG	GCT	GTG	CGG	CTG	GAC	CTC	ATC	3430
Leu	Phe	Thr	Cys	Ala	Met	Arg	Trp	Leu	Ala	Val	Arg	Leu	Asp	Leu	Ile	
1090					1095					1100					1105	
AGC	ATC	GCC	CTC	ATC	ACC	ACC	ACG	GGG	CTG	ATG	ATC	GTT	CTT	ATG	CAC	3478
Ser	Ile	Ala	Leu	Ile	Thr	Thr	Thr	Gly	Leu	Met	Ile	Val	Leu	Met	His	
1110					1115					1120						
GGG	CAG	ATT	CCC	CCA	GCC	TAT	GCG	GGT	CTC	GCC	ATC	TCT	TAT	GCT	GTC	3526
Gly	Gln	Ile	Pro	Pro	Ala	Tyr	Ala	Gly	Leu	Ala	Ile	Ser	Tyr	Ala	Val	
1125					1130					1135						
CAG	TTA	ACG	GGG	CTG	TTC	CAG	TTT	ACG	GTC	AGA	CTG	GCA	TCT	GAG	ACA	3574
Gln	Leu	Thr	Gly	Leu	Phe	Gln	Phe	Thr	Val	Arg	Leu	Ala	Ser	Glu	Thr	
1140					1145					1150						
GAA	GCT	CGA	TTC	ACC	TCG	GTG	GAG	AGG	ATC	AAT	CAC	TAC	ATT	AAG	ACT	3622
Glu	Ala	Arg	Phe	Thr	Ser	Val	Glu	Arg	Ile	Asn	His	Tyr	Ile	Lys	Thr	
1155					1160					1165						
CTG	TCC	TTG	GAA	GCA	CCT	GCC	AGA	ATT	AAG	AAC	AAG	GCT	CCC	TCC	CCT	3670
Leu	Ser	Leu	Glu	Ala	Pro	Ala	Arg	Ile	Lys	Asn	Lys	Ala	Pro	Ser	Pro	
1170					1175					1180					1185	
GAC	TGG	CCC	CAG	GAG	GGA	GAG	GTG	ACC	TTT	GAG	AAC	GCA	GAG	ATG	AGG	3718
Asp	Trp	Pro	Gln	Glu	Gly	Glu	Val	Thr	Phe	Glu	Asn	Ala	Glu	Met	Arg	

B1
Conf

1190

1195

1200

TAC CGA GAA AAC CTC CCT CTC GTC CTA AAG AAA GTA TCC TTC ACG ATC	3766
Tyr Arg Glu Asn Leu Pro Leu Val Leu Lys Lys Val Ser Phe Thr Ile	
1205 1210 1215	
AAA CCT AAA GAG AAG ATT GGC ATT GTG GGG CGG ACA GGA TCA GGG AAG	3814
Lys Pro Lys Glu Lys Ile Gly Ile Val Gly Arg Thr Gly Ser Gly Lys	
1220 1225 1230	
TCC TCG CTG GGG ATG GCC CTC TTC CGT CTG GTG GAG TTA TCT GGA GGC	3862
Ser Ser Leu Gly Met Ala Leu Phe Arg Leu Val Glu Leu Ser Gly Gly	
1235 1240 1245	
TGC ATC AAG ATT GAT GGA GTG AGA ATC AGT GAT ATT GGC CTT GCC GAC	3910
Cys Ile Lys Ile Asp Gly Val Arg Ile Ser Asp Ile Gly Leu Ala Asp	
1250 1255 1260 1265	
CTC CGA AGC AAA CTC TCT ATC ATT CCT CAA GAG CCG GTG CTG TTC AGT	3958
Leu Arg Ser Lys Leu Ser Ile Ile Pro Gln Glu Pro Val Leu Phe Ser	
1270 1275 1280	
GGC ACT GTC AGA TCA AAT TTG GAC CCC TTC AAC CAG TAC ACT GAA GAC	4006
Gly Thr Val Arg Ser Asn Leu Asp Pro Phe Asn Gln Tyr Thr Glu Asp	
1285 1290 1295	
CAG ATT TGG GAT GCC CTG GAG AGG ACA CAC ATG AAA GAA TGT ATT GCT	4054
Gln Ile Trp Asp Ala Leu Glu Arg Thr His Met Lys Glu Cys Ile Ala	
1300 1305 1310	
CAG CTA CCT CTG AAA CTT GAA TCT GAA GTG ATG GAG AAT GGG GAT AAC	4102
Gln Leu Pro Leu Lys Leu Ser Glu Val Met Glu Asn Gly Asp Asn	
1315 1320 1325	
TTC TCA GTG GGG GAA CGG CAG CTC TTG TGC ATA GCT AGA GCC CTG CTC	4150
Phe Ser Val Gly Glu Arg Gln Leu Leu Cys Ile Ala Arg Ala Leu Leu	
1330 1335 1340 1345	
CGC CAC TGT AAG ATT CTG ATT TTA GAT GAA GCC ACA GCT GCC ATG GAC	4198
Arg His Cys Lys Ile Leu Ile Leu Asp Glu Ala Thr Ala Ala Met Asp	
1350 1355 1360	
ACA GAG ACA GAC TTA TTG ATT CAA GAG ACC ATC CGA GAA GCA TTT GCA	4246
Thr Glu Thr Asp Leu Leu Ile Gln Glu Thr Ile Arg Glu Ala Phe Ala	
1365 1370 1375	
GAC TGT ACC ATG CTG ACC ATT GCC CAT CGC CTG CAC ACG GTT CTA GGC	4294
Asp Cys Thr Met Leu Thr Ile Ala His Arg Leu His Thr Val Leu Gly	
1380 1385 1390	
TCC GAT AGG ATT ATG GTG CTG GCC CAG GGA CAG GTG GTG GAG TTT GAC	4342
Ser Asp Arg Ile Met Val Leu Ala Gln Gly Gln Val Val Glu Phe Asp	
1395 1400 1405	
ACC CCA TCG GTC CTT CTG TCC AAC GAC AGT TCC CGA TTC TAT GCC ATG	4390
Thr Pro Ser Val Leu Leu Ser Asn Asp Ser Arg Phe Tyr Ala Met	
1410 1415 1420 1425	
TTT GCT GCT GCA GAG AAC AAG GTC GCT GTC AAG GGC TGA CTCTCC	4436
Phe Ala Ala Ala Glu Asn Lys Val Ala Val Lys Gly	
1430 1435	
CTGTTGACGA AGTCTCTTTT CTTTAGAGCA TTGCCATTCC CTGCCTGGGG CGGGCCCCCTT	4496
CATCGCGTCC TCCTACCGAA ACCTTGCCTT TCTCGATTTT ATCTTTTCGCA CAGCAGTTCC	4556
GGATTGGCTT GTGTGTTTCA CTTTtagGGA GAGTCATATT TTGATTATTG TATTTATTCC	4616
ATATTCATGT AAACAAAATT TAGTTTTTGT TCTTAATTGC ACTCTAAAAG GTTCAGGGAA	4676
CCGTTATTAT AATTGTATCA GAGGCCTATA ATGAAGCTTT ATACGTGTAG CTATATCTAT	4736

B1
Cmt

ATATAATTCT GTACATAGCC TATATTTACA GTGAAATGT AAGCTGTTTA TTTTATATTA 4796
 AAATAAGCAC TGTGCTAAAA AAAAAAAAAA AAAAAAAAAA AGGGCGGCCG C 4847

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1437 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Lys Asp Ile Asp Ile Gly Lys Glu Tyr Ile Ile Pro Ser Pro Gly
 1 5 10 15
 Tyr Arg Ser Val Arg Glu Arg Thr Ser Thr Ser Gly Thr His Arg Asp
 20 25 30
 Arg Glu Asp Ser Lys Phe Arg Arg Thr Arg Pro Leu Glu Cys Gln Asp
 35 40 45
 Ala Leu Glu Thr Ala Ala Arg Ala Glu Gly Leu Ser Leu Asp Ala Ser
 50 55 60
 Met His Ser Gln Leu Arg Ile Leu Asp Glu Glu His Pro Lys Gly Lys
 65 70 75 80
 Tyr His His Gly Leu Ser Ala Leu Lys Pro Ile Arg Thr Thr Cys Lys
 85 90 95
 His Gln His Pro Val Asp Asn Ala Gly Leu Phe Ser Cys Met Thr Phe
 100 105 110
 Ser Trp Leu Ser Ser Leu Ala Arg Val Ala His Lys Lys Gly Glu Leu
 115 120 125
 Ser Met Glu Asp Val Trp Ser Leu Ser Lys His Glu Ser Ser Asp Val
 130 135 140
 Asn Cys Arg Arg Leu Glu Arg Leu Trp Gln Glu Glu Leu Asn Glu Val
 145 150 155 160
 Gly Pro Asp Ala Ala Ser Leu Arg Arg Val Val Trp Ile Phe Cys Arg
 165 170 175
 Thr Arg Leu Ile Leu Ser Ile Val Cys Leu Met Ile Thr Gln Leu Ala
 180 185 190
 Gly Phe Ser Gly Pro Ala Phe Met Val Lys His Leu Leu Glu Tyr Thr
 195 200 205
 Gln Ala Thr Glu Ser Asn Leu Gln Tyr Ser Leu Leu Leu Val Leu Gly
 210 215 220
 Leu Leu Leu Thr Glu Ile Val Arg Ser Trp Ser Leu Ala Leu Thr Trp
 225 230 235 240
 Ala Leu Asn Tyr Arg Thr Gly Val Arg Leu Arg Gly Ala Ile Leu Thr
 245 250 255
 Met Ala Phe Lys Lys Ile Leu Lys Leu Lys Asn Ile Lys Glu Lys Ser
 260 265 270
 Leu Gly Glu Leu Ile Asn Ile Cys Ser Asn Asp Gly Gln Arg Met Phe
 275 280 285
 Glu Ala Ala Ala Val Gly Ser Leu Leu Ala Gly Gly Pro Val Val Ala
 290 295 300

B!
 Cm

Ile Leu Gly Met Ile Tyr Asn Val Ile Ile Leu Gly Pro Thr Gly Phe
 305 310 315 320
 Leu Gly Ser Ala Val Phe Ile Leu Phe Tyr Pro Ala Met Met Phe Ala
 325 330 335
 Ser Arg Leu Thr Ala Tyr Phe Arg Arg Lys Cys Val Ala Ala Thr Asp
 340 345 350
 Glu Arg Val Gln Lys Met Asn Glu Val Leu Thr Tyr Ile Lys Phe Ile
 355 360 365
 Lys Met Tyr Ala Trp Val Lys Ala Phe Ser Gln Ser Val Gln Lys Ile
 370 375 380
 Arg Glu Glu Glu Arg Arg Ile Leu Glu Lys Ala Gly Tyr Phe Gln Ser
 385 390 395 400
 Ile Thr Val Gly Val Ala Pro Ile Val Val Val Ile Ala Ser Val Val
 405 410 415
 Thr Phe Ser Val His Met Thr Leu Gly Phe Asp Leu Thr Ala Ala Gln
 420 425 430
 Ala Phe Thr Val Val Thr Val Phe Asn Ser Met Thr Phe Ala Leu Lys
 435 440 445
 Val Thr Pro Phe Ser Val Lys Ser Leu Ser Glu Ala Ser Val Ala Val
 450 455 460
 Asp Arg Phe Lys Ser Leu Phe Leu Met Glu Glu Val His Met Ile Lys
 465 470 475 480
 Asn Lys Pro Ala Ser Pro His Ile Lys Ile Glu Met Lys Asn Ala Thr
 485 490 495
 Leu Ala Trp Asp Ser Ser His Ser Ser Ile Gln Asn Ser Pro Lys Leu
 500 505 510
 Thr Pro Lys Met Lys Lys Asp Lys Arg Ala Ser Arg Gly Lys Lys Glu
 515 520 525
 Lys Val Arg Gln Leu Gln Arg Thr Glu His Gln Ala Val Leu Ala Glu
 530 535 540
 Gln Lys Gly His Leu Leu Leu Asp Ser Asp Glu Arg Pro Ser Pro Glu
 545 550 555 560
 Glu Glu Glu Gly Lys His Ile His Leu Gly His Leu Arg Leu Gln Arg
 565 570 575
 Thr Leu His Ser Ile Asp Leu Glu Ile Gln Glu Gly Lys Leu Val Gly
 580 585 590
 Ile Cys Gly Ser Val Gly Ser Gly Lys Thr Ser Leu Ile Ser Ala Ile
 595 600 605
 Leu Gly Gln Met Thr Leu Leu Glu Gly Ser Ile Ala Ile Ser Gly Thr
 610 615 620
 Phe Ala Tyr Val Ala Gln Gln Ala Trp Ile Leu Asn Ala Thr Leu Arg
 625 630 635 640
 Asp Asn Ile Leu Phe Gly Lys Glu Tyr Asp Glu Glu Arg Tyr Asn Ser
 645 650 655
 Val Leu Asn Ser Cys Cys Leu Arg Pro Asp Leu Ala Ile Leu Pro Ser
 660 665 670
 Ser Asp Leu Thr Glu Ile Gly Glu Arg Gly Ala Asn Leu Ser Gly Gly
 675 680 685

B
 Cnt

Gln Arg Gln Arg Ile Ser Leu Ala Arg Ala Leu Tyr Ser Asp Arg Ser
 690 695 700
 Ile Tyr Ile Leu Asp Asp Pro Leu Ser Ala Leu Asp Ala His Val Gly
 705 710 715 720
 Asn His Ile Phe Asn Ser Ala Ile Arg Lys His Leu Lys Ser Lys Thr
 725 730 735
 Val Leu Phe Val Thr His Gln Leu Gln Tyr Leu Val Asp Cys Asp Glu
 740 745 750
 Val Ile Phe Met Lys Glu Gly Cys Ile Thr Glu Arg Gly Thr His Glu
 755 760 765
 Glu Leu Met Asn Leu Asn Gly Asp Tyr Ala Thr Ile Phe Asn Asn Leu
 770 775 780
 Leu Leu Gly Glu Thr Pro Pro Val Glu Ile Asn Ser Lys Lys Glu Thr
 785 790 795 800
 Ser Gly Ser Gln Lys Lys Ser Gln Asp Lys Gly Pro Lys Thr Gly Ser
 805 810 815
 Ile Lys Lys Glu Lys Ala Val Lys Pro Glu Glu Gly Gln Leu Val Gln
 820 825 830
 Leu Glu Glu Lys Gly Gln Gly Ser Val Pro Trp Ser Val Tyr Gly Val
 835 840 845
 Tyr Ile Gln Ala Ala Gly Gly Pro Leu Ala Phe Leu Val Ile Met Ala
 850 855 860
 Leu Phe Met Leu Asn Val Gly Ser Thr Ala Phe Ser Thr Trp Trp Leu
 865 870 875 880
 Ser Tyr Trp Ile Lys Gln Gly Ser Gly Asn Thr Thr Val Thr Arg Gly
 885 890 895
 Asn Glu Thr Ser Val Ser Asp Ser Met Lys Asp Asn Pro His Met Gln
 900 905 910
 Tyr Tyr Ala Ser Ile Tyr Ala Leu Ser Met Ala Val Met Leu Ile Leu
 915 920 925
 Lys Ala Ile Arg Gly Val Val Phe Val Lys Gly Thr Leu Arg Ala Ser
 930 935 940
 Ser Arg Leu His Asp Glu Leu Phe Arg Arg Ile Leu Arg Ser Pro Met
 945 950 955 960
 Lys Phe Phe Asp Thr Thr Pro Thr Gly Arg Ile Leu Asn Arg Phe Ser
 965 970 975
 Lys Asp Met Asp Glu Val Asp Val Arg Leu Pro Phe Gln Ala Glu Met
 980 985 990
 Phe Ile Gln Asn Val Ile Leu Val Phe Phe Cys Val Gly Met Ile Ala
 995 1000 1005
 Gly Val Phe Pro Trp Phe Leu Val Ala Val Gly Pro Leu Val Ile Leu
 1010 1015 1020
 Phe Ser Val Leu His Ile Val Ser Arg Val Leu Ile Arg Glu Leu Lys
 1025 1030 1035 1040
 Arg Leu Asp Asn Ile Thr Gln Ser Pro Phe Leu Ser His Ile Thr Ser
 1045 1050 1055
 Ser Ile Gln Gly Leu Ala Thr Ile His Ala Tyr Asn Lys Gly Gln Glu
 1060 1065 1070

B1
 Amf

Phe Leu His Arg Tyr Gln Glu Leu Leu Asp Asp Asn Gln Ala Pro Phe
 1075 1080 1085
 Phe Leu Phe Thr Cys Ala Met Arg Trp Leu Ala Val Arg Leu Asp Leu
 1090 1095 1100
 Ile Ser Ile Ala Leu Ile Thr Thr Thr Gly Leu Met Ile Val Leu Met
 1105 1110 1115 1120
 His Gly Gln Ile Pro Pro Ala Tyr Ala Gly Leu Ala Ile Ser Tyr Ala
 1125 1130 1135
 Val Gln Leu Thr Gly Leu Phe Gln Phe Thr Val Arg Leu Ala Ser Glu
 1140 1145 1150
 Thr Glu Ala Arg Phe Thr Ser Val Glu Arg Ile Asn His Tyr Ile Lys
 1155 1160 1165
 Thr Leu Ser Leu Glu Ala Pro Ala Arg Ile Lys Asn Lys Ala Pro Ser
 1170 1175 1180
 Pro Asp Trp Pro Gln Glu Gly Glu Val Thr Phe Glu Asn Ala Glu Met
 1185 1190 1195 1200
 Arg Tyr Arg Glu Asn Leu Pro Leu Val Leu Lys Lys Val Ser Phe Thr
 1205 1210 1215
 Ile Lys Pro Lys Glu Lys Ile Gly Ile Val Gly Arg Thr Gly Ser Gly
 1220 1225 1230
 Lys Ser Ser Leu Gly Met Ala Leu Phe Arg Leu Val Glu Leu Ser Gly
 1235 1240 1245
 Gly Cys Ile Lys Ile Asp Gly Val Arg Ile Ser Asp Ile Gly Leu Ala
 1250 1255 1260
 Asp Leu Arg Ser Lys Leu Ser Ile Ile Pro Gln Glu Pro Val Leu Phe
 1265 1270 1275 1280
 Ser Gly Thr Val Arg Ser Asn Leu Asp Pro Phe Asn Gln Tyr Thr Glu
 1285 1290 1295
 Asp Gln Ile Trp Asp Ala Leu Glu Arg Thr His Met Lys Glu Cys Ile
 1300 1305 1310
 Ala Gln Leu Pro Leu Lys Leu Glu Ser Glu Val Met Glu Asn Gly Asp
 1315 1320 1325
 Asn Phe Ser Val Gly Glu Arg Gln Leu Leu Cys Ile Ala Arg Ala Leu
 1330 1335 1340
 Leu Arg His Cys Lys Ile Leu Ile Leu Asp Glu Ala Thr Ala Ala Met
 1345 1350 1355 1360
 Asp Thr Glu Thr Asp Leu Leu Ile Gln Glu Thr Ile Arg Glu Ala Phe
 1365 1370 1375
 Ala Asp Cys Thr Met Leu Thr Ile Ala His Arg Leu His Thr Val Leu
 1380 1385 1390
 Gly Ser Asp Arg Ile Met Val Leu Ala Gln Gly Gln Val Val Glu Phe
 1395 1400 1405
 Asp Thr Pro Ser Val Leu Leu Ser Asn Asp Ser Ser Arg Phe Tyr Ala
 1410 1415 1420
 Met Phe Ala Ala Ala Glu Asn Lys Val Ala Val Lys Gly
 1425 1430 1435

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 463 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCGTCCGCCT AGAACGCAGA GATGAGGTAC CGAGAAAACC TCCCTCTCGT CCTAAAGAAA 60
GTATCCTTCA CGATCAAACC TAAAGAGAAG ATTGGCATTG TGGGGCGGAC AGGATCAGGG 120
AAGTCCTCGC TGGGGATGGC CCTCTCCGT CTGGTGGAGT TATCTGGAGG CTGCATCAAG 180
ATTGATGGAG TGAGAATCAG TGATATTGGC CTTGCCGACC TCCGAAGCAA ACTCTCTATC 240
ATTCCTCAAG AGCCGGTGCT GTTCAGTGGC ACTGTCAGAT CAAATTTGGA CCCTTCAACC 300
AGTACACTGA AGACCAGATT TGGGATGCCC TGGAAAGGAC ACACATGAAA GAATGTATTG 360
CTCCAGCTAC CTCCTGAAAC TTGAATCCTG AATTTGATGG AGAAATGGGG AAATAACTTC 420
TCCAGTTGGG GGAAACGGCA CTCTTTGTTG CCATACCTAN ACC 463

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TGCTGGTTCT CTCCCTCACA CTTC 24

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTCGGCTCGG GCTGCTGTTT CCAA 24

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGTGCTGGT GTTTGGAAGT AGTC 24

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCCAGAGAAG AAAGCCACGA AAAA

24

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGCACACGA TGGACAGGAT GAGC

24

S'
Gry